BS

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/709,659	11/10/2000	Adam Roth	4056-4000	4294

7590 07/15/2004

Morgan & Finnegan LLP
345 Park Avenue
New York, NY 10154

EXAMINER	
NOLAN, DANIEL A	
ART UNIT	**PAPER NUMBER**
2654	13

DATE MAILED: 07/15/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/709,659	ROTH ET AL.	
	Examiner	Art Unit	
	Daniel A. Nolan	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *04 March 2004* .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-15,30-44,59-73,88,90 and 92* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-15,30-44,59-73,88,90 and 92* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on *08 December 2003* is: a)☒ approved b)☐ disapproved by the Examiner

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Response to Amendment

2. The response filed 04 March 2004 was entered with the following effect:

- The claims were changed as indicated and examined on the merits.

Claim Rejections - 35 USC § 103

3. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

Johnson *et al*[(s)], Cosatto *et al* & Kawamoto

4. Claims 1-5, 30-34 and 59-63 are rejected under 35 U.S.C. 103(a) as being unpatentable over the combined references of Johnson *et al*['383] (U.S. Patent 5,568,383) in view of Johnson *et al*['910] (U.S. Patent 5,434,910) and further in view of Cosatto *et al* (U.S. Patent 6,112,177) and further in view of Kawamoto (U.S. Patent 6,169,902 B1).

5. Regarding claims 1, 30 and 59, the Background provided by Johnson *et al*['383] for the Document Transmission Network portion of their invention teaches *operating in a network* (the LAN of figure 9 and the *network* of claim 7, lines 19-21) and adding multimedia to E-Mail both to send documents other than text and to enhance the E-Mail aesthetically with the description of underlying art provided in the Background (column 1 lines 21-46). Johnson *et al*['383] does not specifically mention using these capabilities to embellish mailed documents with an avatar in the form of a likeness of a sender.

With the invention for Co-Articulation for Audio-Visual Text-To-Speech Synthesis, Cosatto *et al* teaches that it is advantageous to accompany messages with a realistic likeness of the sender, as Background (column 1 lines 19-27). It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Cosatto *et al* to the device/method of Johnson *et al*['383] so as to provide a credible enhancement conveying emotions that will not be misinterpreted.

- In this, Johnson *et al*'383 (column 1 line 26) reads on the feature in the preamble, *for preparing a Multi-Mail message for transmission over a network.*
- Johnson *et al*'383 (column 1 line 23) reads on the feature of *receiving data comprising textual content of said message;*
- Cosatto *et al* (column 1 lines 30-32) read on the features of *creating one or more multimedia components associated with said message,* where the *multimedia component represents a likeness of a sender* (as in column 10-13)*; and synthesizing multimedia components with said textual content* (column 3 lines 28-30).
- Regarding the distinct features of claim 59, Johnson *et al*'383 does not specifically mention a *database,* but Johnson *et al*'910 teaches that the *message server* performs the equivalent functions (column 2 lines 43-50) of *storage and retrieval* (column 3 lines 40-51).

  Similarly, Johnson *et al*'383 is silent on the subject that *the CPU is configured for speech processing,* while Johnson *et al*'910 depicts all components of the processor in a manner that agrees with this configuration (304-308 in figure 3).

  These cited combinations would have been expected to be within the experience of and therefore obvious to a person of ordinary skill in the art of speech processing at the time of the invention in order to apply the teachings of Johnson *et al*'910 to the device/method of Johnson *et al*'383 to enable access when required, other than at the time the message is generated.

Neither Johnson et al(s), Cosatto et al nor Kawamoto speak to verifying sender ID, access to sender(s) multimedia and a likeness of the sender. Kawamoto, with the inventions for *information terminal, processing method by information terminal, information providing apparatus and information network system* reads on the features of *verifying identifier information of the sender of the multi-media message* (steps S4 to S8 in figure 6, where S8 is electronic mail, see column 3 lines 8-9); *allowing access to stored multimedia information of the sender* (column 7 line 48 to column 8 line 12), containing a *likeness of the sender of the multi-mail message based on stored multimedia information of the sender* (image F in figure 11 when the map is sent by e-mail).

It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Kawamoto to the device/method of Johnson et al(s) and/or Cosatto et al so that e-mail correspondents might recognize each other in person.

6. Regarding claims 2, 31 and 60; the claims are set forth with the same limits as claims 1, 30 and 59, respectively. Johnson *et al'383* (column 1 line 24) read on the feature that *the multimedia component comprises audio information.*

7. Regarding claims 3, 32 and 61; the claims are set forth with the same limits as claims 1, 30 and 59, respectively. Johnson *et al'383* (with *graphic* in column 1 line 23) read on the feature that *the multimedia component comprises image information.*

8. Regarding claims 4, 33 and 62; the claims are set forth with the same limits as claims 3, 32 and 61, respectively. Johnson *et al*'383 (with the active nature of text-to-speech described as being *Audient* – *adj. listening: paying attention [Ch]* – in column 1 line 45) read on the feature that *the image information may be static or dynamic.*

9. Regarding claims 5, 34 and 63; the claims are set forth with the same limits as claims 2, 31 and 60, respectively. Johnson *et al*'383 does teach using TTS to generate speech for multimedia in mail messages (column 1 line 45) but does not specifically mention that the voice would be that of the author.

Cosatto *et al* (column 3 line 26-27) uses the author(s) voice to produce TTS speech, reading on the feature that *the audio component comprises voice data that enables the generation of sounds similar to the user(s) voice speaking the words of the textual content of the message.* It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Cosatto *et al* to the device/method of Johnson *et al*'383 so as to not diminish the credibility of the image.

Johnson *et al*(s), Cosatto *et al* & Lee *et al*

10. Claims 6-9, 35-38 and 64-67 are rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson *et al*'383 in view of Johnson *et al*'910 and further in view of

Cosatto *et al* and further in view of Kawamoto and further in view of Lee *et al* (U.S. Patent 6,088,673).

11. With regard to claims 6, 35 and 64; the claims are set forth with the same limits as claims 2, 31 and 60, respectively. While it is common practice in the art of speech synthesis to produce voice using non-specific models to avoid training, economize and make synthetic speech products more flexible, Johnson *et al*'383 is silent on this subject.

The TTS Conversion System for Multimedia invention of Lee *et al* teaches in the Background (column 1 lines 49-55) the feature of *voice data that enables the generation of sounds similar to a generic voice sample*, and subsequently implements (in their Claim 1 lines 57-62) the capability to change voices to match multimedia.

This would have made it obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Lee *et al* to the device/method of Johnson *et al*(s) Cosatto *et al* & Kawamoto when the need arose to synchronize the speech with the multimedia presentation independently of the text.

12. Regarding claims 7, 36 and 65; the claims are set forth with the same limits as claims 2, 31 and 60, respectively. Johnson *et al*'383 reads on the feature that *comprises voice data that enables the generation of any stored sound* (with *music* and sounds in column 1 lines 26-30).

13. With regard to claims 8, 37 and 66; the claims are set forth with the same limits as claims 2, 31 and 60, respectively. Where Johnson *et al*'383 does not mention *parsing,* Cosatto *et al* provides the means in the central processor but is silent on the subject of *segmenting into sentences*, Lee *et al* discloses that the following processing steps will permit synchronization between text and speech:

- Lee *et al* (column 1 lines 62-65) reads on the feature of *parsing the audio information into sentences and* (by specifying the *accent* in column 1 line 42) *for voice modulation controls.*
- Lee *et al* (with the detail in column 1 lines 44-48) teaches the feature of *assigning voice modulation to audio information;*
- Lee *et al* (with the coding at the 1st line in table 1, column 3 line 61), reads on the feature of *sequencing phoneme and modulation information* (at the 5th line)*;*
- Lee *et al* (in column 6 lines 10-65) provides three methods (starting at lines 20, 41 & 53) that read on the feature of *translating said phoneme sequence into a sound component sequence.*

It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Lee *et al* on (*parsing into sentences, modulating voice, sequencing and translating into sound*) to the device/method of Johnson *et al*(s), Cosatto *et al* and/or Kawamoto to synchronize TTS with text for high quality.

14. Regarding claims 9 as understood by the Examiner and claim 67; the claims are set forth with the same limits as claims 1 and 60, respectively. Johnson *et al*'383 does not mention *the synthesis of the image multimedia component,* but Lee *et al* teaches the advantages of combining speech segments with images (column 6 lines 30-33) makes animated synthesis possible.

- Lee *et al* (column 6 lines 20-57 & 63-65) reads on the feature of *identifying speech movement image feature; and*
- Lee *et al* (column 8 lines 37-40) reads on the feature of *generating frames representing movement of said image features.*

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to further modify Johnson *et al*(s), Cosatto *et al* and/or Kawamoto in view of Lee *et al*, such that Johnson *et al*'383 *produces mouth movement*, in order to receive the benefit of an articulate facial image that appears to speak.

Johnson *et al*(s), Cosatto & Kawamoto

15. Claims 10, 39 and 68 are rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson *et al*'383 in view of Johnson *et al*'910 and further in view of Cosatto *et al* and further in view of Kawamoto.

16. Regarding claims 10, 39 and 68; the claims are set forth with the same limits as claims 1, 30 and 59, respectively.

- Johnson *et al*'383 (column 1 line 24) reads on the feature that *the multimedia component comprises audio information,* and
- Johnson *et al*'383 (with *graphic* in column 1 line 23) read on the feature that *the multimedia component comprises image information.*

Johnson *et al*(s), Cosatto *et al*, Kawamoto & Lee *et al*

17. Claims 11-12, 40-41 and 69-70 are rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson *et al*'383 in view of Johnson *et al*'910 and further in view of Cosatto *et al* and further in view of Kawamoto and further in view of Lee *et al*.

18. Regarding claim 38 as understood by the Examiner, the claim is set forth with the same limits as claim 39. The features of the claim are the same as those found in claim 9, above, and the claim is rejected for the same reasons addressed in response.

19. Regarding claims 11, 40 and 69; the claims are set forth with the same limits as claims 10, 39 and 68, respectively. Johnson *et al*'383 is silent on the *sequencing* issue. Lee *et al* teaches arranging phonemes, mouth frame time and speech movement to correspond to articulate speech.

- Lee *et al* (with the coding at the 1st line in table 1, column 3 line 61), reads on the feature of *composing a phoneme sequence;*

- Lee *et al* (column 4 line 65 and column 5 lines 66-67) reads on the feature of *composing a mouth frame time sequence which matches the phoneme time sequence;*
- Lee *et al* (column 6 lines 20-30) reads on the feature of composing *speech movement image frame sequence; and*
- Lee *et al* (column 8 lines 28-34) reads on the feature of *combining the image and phoneme sequences.*

This would have made it obvious to one of ordinary skill in the art at the time of the invention to modify Johnson *et al*[(s)], Cosatto *et al* and/or Kawamoto in view of Lee *et al*, such that Johnson *et al*['383] includes *composing phoneme with matching images of mouth frame time* into *speech movement image frame sequence* to produce *combined image and phoneme sequences* in order that the image would credibly appear to recite the same text.

20. Regarding claims 12, 41 and 70; the claims are set forth with the same limits as claims 10, 39 and 68, respectively. Johnson *et al*[(s)], Cosatto *et al* and Kawamoto are silent on the subject of *varying components.* Lee *et al* teaches diversifying speech (in column 7 lines 10-12) by *varying one or more of said components to convey one or more senses of said message content,* which would have made it obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Lee *et al* to the device/method of Johnson *et al*[(s)], Cosatto *et al*

and/or Kawamoto so as to alter the synthesized text, for example, by the familiar practice of simulating with a different voice or *sotto voce* when quoting or speaking an aside.

Johnson *et al*'383, Cosatto *et al*, Kawamoto Lee *et al* & Kirksey *et al*

21. Claims 13, 42 and 71 are rejected under 35 U.S.C. 103(a) as being rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson *et al*'383 in view of Johnson *et al*'910 and further in view of Cosatto *et al* and further in view of Lee *et al* and further in view of Kirksey *et al* (U.S. Patent 5,938,447 A).

22. Regarding claims 13, 42 and 71; the claims are set forth with the same limits as claims 12, 41 and 70, respectively. Johnson *et al*'383 is silent on the subject of emotions in multimedia or e-mail. Kirksey *et al* teaches the ability for making an audio-visual work with a series of visual word symbols coordinated with oral word utterances to convey additional meaning in the form of emotions, with the table (column 11 lines 25-32) reading on the feature that *the senses of said message content correspond to one or more sender emotions associated with said message.* It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Kirksey *et al* to the device/method of Johnson *et al*(s), Cosatto *et al* and/or Kawamoto so as to have the appearance of a text match the context of the message.

Johnson *et al*[(s)], Cosatto *et al*, Kawamoto, Lee *et al*, Kirksey *et al* & Skelly

23. Claims 14-15, 43-44 and 72-73 are rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson *et al*['383] in view of Johnson *et al*['910] and further in view of Cosatto *et al* and further in view of Kawamoto and further in view of Lee *et al* and further in view of Kirksey *et al* and further in view of Skelly (U.S. Patent 6,064,383).

24. Regarding claims 14, 43 and 72; the claims are set forth with the same limits as claims 13, 42 and 71, respectively. Johnson *et al*['383] discloses *image components* but does not does not disclose *emotions* with relation to *image.* Skelly teaches *selecting an emotional appearance for a graphical character* that discloses the feature that *the sender emotions are conveyed by a manipulating one or more said image components* (64-66 in figure 5). It would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Skelly to the device/method of Johnson *et al*[(s)], Cosatto *et al* and/or Kawamoto so as to realize the benefit of displaying emotion beyond what words can convey.

25. Regarding claims 15, 44 and 73; the claims are set forth with the same limits as claims 13, 42 and 71, respectively. Johnson *et al*['383] discloses *audio components* but does not does not disclose *emotions* with relation to *audio.* Skelly teaches *selecting an emotional prosody for a graphical character* that discloses the feature that *the sender emotions are conveyed by a manipulating one or more said audio components* (claims 3 & 7, column 8 lines 22-23 & 40-43). It would have been obvious to a person of

ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Skelly to the device/method of Johnson *et al*[(s)], Cosatto *et al* and/or Kawamoto so as to realize the benefit of displaying emotion beyond what words can convey.

Johnson *et al*[(s)], Cosatto *et al*, Kawamoto & Lee *et al*

26. Claims 88, 90 and 92 are rejected under 35 U.S.C. 103(a) as being unpatentable over Johnson *et al*['383] in view of Johnson *et al*['910] and further in view of Cosatto *et al* and further in view of Kawamoto and further in view of Lee *et al*.

27. Regarding claims 88, 90 and 92, Johnson *et al*[(s)] are silent on the matter of *code*. Lee *et al* (in tables 1 and 2, columns 3-5) provides at least a pseudocode that will *prepare a Multi-Mail message*. Indicating the need for this feature, Johnson *et al*['383] describes the problem of words becoming "lost in translation" and in doing so, makes it obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Lee *et al* to the device/method of Johnson *et al*['383] thus providing a description to avoid the expense and uncertainty of translation.

Regarding the other features of the claims:

- Johnson *et al*['383] (column 1 lines 24-26) reads on the features contained in the preamble of the claims as that the *message* (is) *for transmission over a network.*

- With respect to claims 88 and 92, where Johnson *et al*'383 is silent as to the issue of *medium* for both claims, Cosatto *et al* shows the intimate arrangement of processor operations and libraries (11 & 14 in figure 2) and, with the disclosure that *the processor consults circuitry or software* (column 10 lines 27-28) reads on the feature that the *software code is on computer readable medium* which would have made it obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method/teachings of Cosatto *et al* to the device/method of Johnson *et al*'383 so as to employ the elements of a conventional PC.
- With respect to the *memory* of claim 90, Johnson *et al*'383 is silent on the subject of processing allocations. Cosatto *et al* discloses (column 5 lines 38-39) that code is in memory and (column 8 lines 45-49 and 55-67) shows the advantages of performing operations in memory and so reads on the feature of *a memory having at least one region for storing computer executable program code.*

  With Cosatto *et al* teaching the desirability of processing while the subject is speaking, it would have been obvious to a person of ordinary skill in the art of speech signal processing at the time of the invention to apply the method-teachings of Cosatto *et al* to the device/method of Johnson *et al*'383 and implement the faster memory of a PC over the relatively slower storage.
- The remaining features of the claims to receive, *create, and synthesize* are the same as those found in claims 1, 30 and 59 and the claims are rejected for the same reasons.

Conclusion

28. The prior art made of record and not relied upon is considered pertinent to applicant(s) disclosure.

- Vaudreuil (U.S. Patent 5,621,727 A) for private addressing plans using community addressing.
- Guillemin (U.S. Patent 6,751,589 B1) voice-actuated generation of documents containing photographic identification.

29. Any inquiry concerning this communication or earlier communications from the Examiner should be directed to Daniel A. Nolan at telephone (703) 305-1368 whose normal business hours are Mon, Tue, Thu & Fri, from 7 AM to 5 PM.

If attempts to contact the examiner by telephone are unsuccessful, supervisor Richemond Dorvil can be reached at (703)305-9645.

The fax phone number for Technology Center 2600 is (703)872-9314. Label informal and draft communications as "DRAFT" or "PROPOSED", & designate formal communications as "EXPEDITED PROCEDURE". Formal response to this action may be faxed according to the above instructions,

or mailed to:

P.O. Box 1450
Alexandria, VA 22313-1450

or hand-deliver to: Crystal Park 2,
2121 Crystal Drive, Arlington, VA,
Sixth Floor (Receptionist).

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to Technology Center 2600 Customer Service Office at telephone number (703) 306-0377.

Daniel A. Nolan
Examiner
Art Unit 2654

DAN/d
July 5, 2004

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER